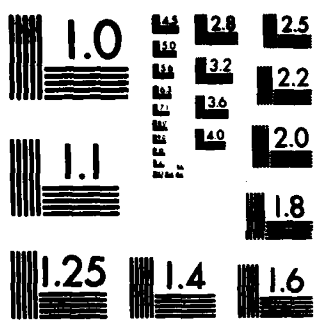
AD-A129 711 STUDY OF ALL-MICROPROCESSOR COMPUTER COMMUNICATION NETWORKS(U) COLUMBIA UNIV NEW YORK DEPT OF ELECTRICAL ENGINEERING AND COM.. M SCHWARTZ ET AL. 27 MAY 83 1/1

UNCLASSIFIED ARO-14292.4-EL DAAG29-79-C-0196 F/G 5/1 NL

END
DATE FILMED
7-83
DTIC


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 14292.4-EL; 16619.4-EL	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) STUDY OF ALL-MICROPROCESSOR COMPUTER COMMUNICATION NETWORKS		5. TYPE OF REPORT & PERIOD COVERED final 16 Aug 76 - 31 Mar 83
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Mischa Schwartz, Theodore Bashkow		8. CONTRACT OR GRANT NUMBER(s) DAAG-29-76-G-0322, DAAG-29-79-C-0196
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Electrical Engineering, Columbia University, New York, New York 10027		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office, Post Office Box 12211, Research Triangle Park, NC 27709		12. REPORT DATE 5/27/83
		13. NUMBER OF PAGES 9
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

DTIC ELECTE JUN 22 1983 S H D

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

communication networks
computer networks
microprocessors
distributed processing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The overall problem studied was that of the organization and utilization of very large microprocessor-based computer communication networks. The first phase of this research was directed toward the actual design of microprocessor based communication networks. It was concerned with the manner in which work should be organized and in the internal structure of the physical nodes of the networks. In the second phase it was assumed (given the first phase results) that such distributed networks, containing very large numbers of nodes, could indeed be constructed. The focus of attention was now on the problem of routing, and distributed processing in large, possibly dynamically varying networks.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

STUDY OF ALL-MICROPROCESSOR COMPUTER COMMUNICATION NETWORKS

FINAL REPORT

May 27, 1983

U. S. ARMY RESEARCH OFFICE

DAAG29-76-G-0322

and

DAAG29-79-C-0196

Columbia University

Department of Electrical Engineering

Seeley W. Mudd Building
New York, New York 10027

Approved for Public Release; Distribution unlimited.

This report covers work carried out under two ARO grants, DAAG29-76-G-0322 and DAAG29-79-C-0196, the latter a follow-on to the first. Work under the first grant will be referred to as the first phase of research, work under the second grant will be referred to as the second phase.

The overall problem studied was that of the organization and utilization of very large microprocessor-based computer communication networks.

The first phase of this research was directed toward the actual design of microprocessor based communication networks. It was concerned with the manner in which work should be organized and in the internal structure of the physical nodes of the networks.

In the second phase it was assumed (given the first phase results) that such distributed networks, containing very large numbers of nodes, could indeed be constructed. The focus of attention was now on the problem of routing, and distributed processing in large, possibly dynamically varying networks.

In phase 1 various applications were modeled in terms of foreground and background tasks. Foreground tasks were those supporting user-terminal interactions at a computing node. Background tasks performed the data processing (or other computations) plus the network communication functions.

Queueing models were used to investigate the ways in which tasks should be assigned to processors[1,2] and on the performance which could be expected in a realistic multiple processor system. This work, begun in phase one, was carried through to completion in phase two.

The foreground tasks are generally smaller but more time-critical than the background tasks. A foreground task, on completion, may generate a background task, and vice-versa. Time-slicing of the background tasks was also incorporated in the models that were developed.

The exact analysis of a multiprocessor system operating under a preemptive

priority scheduling discipline was made. The analysis is new and extended some related work previously reported in the literature. A new approximate procedure was described to analyze a multiprocessor system using a nonpreemptive priority discipline. Under this procedure, the complexity of the problem is first reduced by lumping states, and then generating functions are utilized to derive the required performance measures.

The multiprocessor models analyzed were shown to be suitable for many real applications, such as a time sharing system or a "node" of a distributed processing system, where one or more arrivals to the system trigger some background processing which, on completion, may send the results (response) back to the initiator. These models are made more realistic by the incorporation of task startup and state saving overhead, and contention at the shared memory. A procedure to estimate shared memory interference in a simple but realistic way was developed by extending some analysis reported in the literature. Several different multiple processor configurations were compared with each other using parameter values that are typical for some real applications.

Some portion of the work was devoted to the modeling of systems where task spawning does not take place. One such two-processor, two-class system (using nonpreemptive priorities) was analyzed (approximately) to obtain explicit expressions for the average waiting times. A general multiprocessor system (using nonpreemptive priorities) having an arbitrary number of classes and generally distributed service times was also analyzed using a simplistic approach, but this aproach causes relatively large errors in the results[3]. A number of papers are being prepared on this work. One has already been submitted for publication[4]. Details appear in a doctoral dissertation completed[3].

Study of a desirable internal nodal structure led to multiprocessor design in which one microprocessor handled the data processing load and a second microprocessor dealt with the communication load[5,6].

In Phase 2 a study was made of the routing function in large networks, or networks characterized by frequent topological changes, or both. Both distributed and hierarchical routing were studied in this work.

In the case of distributed routing two parallel studies were carried out. In one study two simple distributed shortest-path routing algorithms were compared on the basis of convergence time and control packets required to be transmitted during the period of convergence. One algorithm was of the class first introduced in the ARPA network. The second was a modified version designed to reduce the number of control packets generated. Both analyses and computer experimentation were used in the comparison[7].

In the second study two new event driven distributed route table update algorithms, A and B, were introduced and proven correct. Algorithm A requires less buffer space to store route tables than other event driven algorithms. Algorithm B, a variation of algorithm A, allows each node to maintain a source tree, i.e. a tree rooted at the node and containing the shortest path to all possible destinations in the network. The source tree may be used to implement source routing, i.e. the whole path from source to destination is determined at the source. Transient route table looping was also studied for algorithms A and B, as well as other algorithms in the literature. Two papers on this material have been presented[8,9].

Hierarchical routing has been suggested in the literature as a means to reduce the size of the route tables when networks become very large. Reduction of the table size may also reduce the communication cost incurred during the update of the route tables. A possible effect of the shrinking of the route tables is that the resulting paths are not optimal. A classification of hierarchical routing schemes was introduced. The trade-off between route table reduction and path length increase was studied in detail for two classes of schemes. Alternate policies for routing in the absence of necessary information were suggested and evaluated.

In order to implement hierarchical routing it is necessary to partition the network into clusters. The network partitioning problem was abstracted to a graph partitioning problem which was shown to be NP complete. A new heuristic procedure, V3.2 was developed which was compared to the agglomerative method, a procedure suggested in the literature. V3.2 was shown to perform considerably better computationally as well as in terms of desired properties of the partitions. The comparison was performed by simulation experiments. An algorithm was developed to randomly generate connected networks suitable to be used in the simulations.

A paper covering the material on hierarchical routing has been submitted for publication, [10]. A doctoral thesis covering the two new algorithms A and B, as well as the work on hierarchical routing has been completed, [11].

A study was also made in Phase 2 of the scheduling of jobs in a distributed computer system.

This problem differs from the classical computer scheduling problem because the entire scheduling problem is not known at any point and neither does the scheduling processor have complete knowledge of the system state. Both of these differences result from the distributed nature of the computer system. Because of these constraints it is not possible to use classical optimal scheduling theory.

Because of the need to compare the performance of a distributed scheduling heuristic with optimal and non-optimal centralized scheduling algorithms a new nonrelative bound for all list scheduling was developed. Using this new bound, the performance distribution of random list schedules was then studied. As far as we know, this is the first such study that has been made.

Classical centralized scheduling theory was then extended to distributed scheduling theory and the performance distribution of a simple network scheduling heuristic was studied. The difference between a centralized algorithm and a distributed algorithm was discussed. The distributed scheduling problem was

presented and the reasons requiring a distributed algorithm were discussed.

Finally, the effect of network delay on scheduling performance was studied and centralized and network scheduling performance were compared. This work on distributed scheduling appears in a doctoral dissertation just completed[12].

A study was also begun in Phase 2 of the problem of distributing resources in a large network. In particular it was desired to optimally allocate copies of a network directory to nodes in the network. Previous models unfortunately fall in the difficult class of an NP-hard problem. However, by limiting the networks to ones having a particular regular topology it is possible to find such an optimal solution which does not "explode" computationally. This work is still underway and has not yet been published. The work turns out to be related to the theory of error correcting codes. As a byproduct of this research, some very useful properties of so-called binary constant weight codes have been found. These codes are nonlinear and very little is known about them.

REFERENCES

1. M. Schwartz, "Throughput and Time Delay Analysis for a Common Queue Configuration in a Multiprocessor Environment," IEEE Transactions on Computers, December 1979.
2. M. Schwartz, "Topological Considerations in Microcomputer Communication Networks: Time Response Analysis," International Conference on Communications, Toronto, 1978.
3. R. Bhatia, "Performance Analysis of Multiple Processor Systems under Priority Scheduling Disciplines," Doctor of Engineering Sciences Dissertation, Columbia University, 1982.
4. R. Bhatia and and M. Schwartz, "A General Shared Memory Interference Model Assuming Constant Memory Access Times," submitted to IEEE Trans. on Computers.
5. T.R. Bashkow, M. Schwartz, P. Pungaliya, R. Karachiwala, J. Hagouel and H. Sullivan, "Node Design for a Distributed Terminal Network," Trends & Applications 1978: Distributed Processing, May 9, 1978, National Bureau of Standards, Gaithersburg, Maryland.
6. T.R. Bashkow, M. Schwartz, P. Pungaliya, R. Karachiwala, J. Hagouel and H. Sullivan, "A Microcomputer Based Communication Network," U.S. Army Workshop on Microprocessors in DoD, Pingree Park Conference Center, Colorado, 22-24 August, 1979.
7. M. Schwartz and Tak-Kim Yum, "Distributed Routing in Computer Communication Networks," Proc. 21st Conference on Decision and Control, Orlando, Fl. Dec. 1982.
8. J. Hagouel and M. Schwartz, "A Distributed Failsafe Route Table Update Algorithm," 3rd International Conference on Distributed Computing Systems, Miami, Fl. Oct. 1982.

9. J. Hagouel, "Source Routing and a Distributed Algorithm to Implement It," IEEE Infocom, San Diego, April 1983.

10. J. Hagouel and M. Schwartz, "Hierarchical Routing in Computer Networks," submitted to IEEE Trans. on Communications.

11. J. Hagouel, "Issues in Routing for Large and Dynamic Networks," doctoral dissertation, Columbia University, May 1983.

12. L.F. Horney, II, "Job Scheduling in a Distributed System," Doctor of Engineering Science Dissertation, Columbia University, 1983.

Degrees Awarded

1. Jack Hagouel, Ph.D
2. Rajiv Bhatia, D. Eng. Sc.
3. Fred Horney, D. Eng. Sc.

Students Supported

1. Jack Hagouel
2. James Kurose
3. Arif Ghafoor
4. Tak-Kim Yum
5. Hamid K. Ahmad
6. Sumitra Ganguly

Papers presented or submitted for publication

1. M. Schwartz, "Throughput and Time Delay Analysis for a Common Queue Configuration in a Multiprocessor Environment," IEEE Transactions on Computers, December 1979.
2. M. Schwartz, "Topological Considerations in Microcomputer Communication Networks: Time Response Analysis," International Conference on Communications, Toronto, 1978.
3. T.R. Bashkow, M. Schwartz, P. Pungaliya, R. Karachiwala, J. Hagouel and H. Sullivan, "Node Design for a Distributed Terminal Network," Trends & Applications 1978: Distributed Processing, May 9, 1978, National Bureau of Standards, Gaithersburg, Maryland.
4. T.R. Bashkow, M. Schwartz, P. Pungaliya, R. Karachiwala, J. Hagouel and H. Sullivan, "A Microcomputer Based Communication Network," U.S. Army Workshop on Microprocessors in DoD, Pingree Park Conference Center, Colorado, 22-24 August, 1979.
5. M. Schwartz and Tak-Kim Yum, "Distributed Routing in Computer Communication Networks," Proc. 21st Conference on Decision and Control, Orlando, Fl. Dec. 1982.

6. J. Hagouel and M. Schwartz, "A Distributed Failsafe Route Table Update Algorithm," 3rd International Conference on Distributed Computing Systems, Miami, Fl. Oct. 1982.

7. J. Hagouel, "Source Routing and a Distributed Algorithm to Implement It," IEEE Infocom, San Diego, April 1983.

8. J. Hagouel and M. Schwartz, "Hierarchical Routing in Computer Networks," submitted to IEEE Trans. on Communications.

9. R. Bhatia and M. Schwartz, "A General Shared Memory Interference Model Assuming Constant Memory Access Times," submitted to IEEE Trans. on Computers.